

WA #5 PD  
PCT  
11-7-97

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,474

DATE: 07/05/2001

TIME: 16:16:43

Input Set : A:\kim.txt

Output Set: N:\CRF3\07032001\I868474.raw

3 <110> APPLICANT: KIM, Jin-Woo  
5 <120> TITLE OF INVENTION: HUMAN CERVICAL CANCER 1 PROTOONCOGENE AND PROTEIN ENCODED  
THEREIN

7 <130> FILE REFERENCE: KIM  
C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/868,474  
C--> 9 <141> CURRENT FILING DATE: 2001-06-15

ENTERED

9 <150> PRIOR APPLICATION NUMBER: KR 1999-44811  
10 <151> PRIOR FILING DATE: 1999-10-15  
12 <150> PRIOR APPLICATION NUMBER: PCT/KR00/00284  
13 <151> PRIOR FILING DATE: 2000-03-30  
15 <160> NUMBER OF SEQ ID NOS: 7  
17 <170> SOFTWARE: PatentIn version 3.0

19 <210> SEQ ID NO: 1  
20 <211> LENGTH: 2118  
21 <212> TYPE: DNA  
22 <213> ORGANISM: Homo sapiens  
24 <220> FEATURE:  
25 <221> NAME/KEY: CDS  
26 <222> LOCATION: (9)..(1088)  
28 <220> FEATURE:  
29 <221> NAME/KEY: sig\_peptide  
30 <222> LOCATION: (9)..(83)  
32 <220> FEATURE:  
33 <221> NAME/KEY: misc\_feature  
34 <222> LOCATION: (435)..(494)  
35 <223> OTHER INFORMATION: transmembrane domain  
38 <400> SEQUENCE: 1

39 ctgtgaag atg gcg ctc tcc agg gtg tgc tgg gct cgg tcg gct gtg tgg 50  
40 Met Ala Leu Ser Arg Val Cys Trp Ala Arg Ser Ala Val Trp  
41 1 5 10  
43 ggc tcg gca gtc acc cct gga cat ttt gtc acc cgg agg ctg caa ctt 98  
44 Gly Ser Ala Val Thr Pro Gly His Phe Val Thr Arg Arg Leu Gln Leu  
45 15 20 25 30  
47 ggt cgc tct ggc ctg gct tgg ggg gcc cct cgg tct tca aag ctt cac 146  
48 Gly Arg Ser Gly Leu Ala Trp Gly Ala Pro Arg Ser Ser Lys Leu His  
49 35 40 45  
51 ctt tct cca aag gca gat gtg aag aac ttg atg tct tat gtg gta acc 194  
52 Leu Ser Pro Lys Ala Asp Val Lys Asn Leu Met Ser Tyr Val Val Thr  
53 50 55 60  
55 aag aca aaa gcg att aat ggg aaa tac cat cgt ttc ttg ggt cgt cat 242  
56 Lys Thr Lys Ala Ile Asn Gly Lys Tyr His Arg Phe Leu Gly Arg His  
57 65 70 75  
59 ttc ccc cgc ttc tat atc ctg tac aca atc ttc atg aaa gga ttg cag 290  
60 Phe Pro Arg Phe Tyr Ile Leu Tyr Thr Ile Phe Met Lys Gly Leu Gln  
61 80 85 90  
63 atg tta tgg gct gat gcc aaa aag gct aga aga ata aag aca aat atg 338  
64 Met Leu Trp Ala Asp Ala Lys Lys Ala Arg Arg Ile Lys Thr Asn Met  
65 95 100 105 110



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,474

DATE: 07/05/2001

TIME: 16:16:43

Input Set : A:\kim.txt

Output Set: N:\CRF3\07032001\I868474.raw

67	tgg aag cac aat ata aag ttt cat caa ctt cca tac cgg gag atg gag	386
68	Trp Lys His Asn Ile Lys Phe His Gln Leu Pro Tyr Arg Glu Met Glu	
69	115 120 125	
71	cat ttg aga cag ttc cgc caa gac gtc acc aag tgt ctt ttc cta ggt	434
72	His Leu Arg Gln Phe Arg Gln Asp Val Thr Lys Cys Leu Phe Leu Gly	
73	130 135 140	
75	att att tcc att cca cct ttt gcc aac tac ctg gtc ttc ttg cta atg	482
76	Ile Ile Ser Ile Pro Pro Phe Ala Asn Tyr Leu Val Phe Leu Leu Met	
77	145 150 155	
79	tac ctg ttt ccc agg caa cta ctg atc agg cat ttc tgg acc cca aaa	530
80	Tyr Leu Phe Pro Arg Gln Leu Leu Ile Arg His Phe Trp Thr Pro Lys	
81	160 165 170	
83	caa caa act gat ttc tta gat atc tat cat gct ttc cgg aag cag tcc	578
84	Gln Gln Thr Asp Phe Leu Asp Ile Tyr His Ala Phe Arg Lys Gln Ser	
85	175 180 185 190	
87	cac cca gaa att att agt tat tta gaa aag gtc atc cct ctc att tct	626
88	His Pro Glu Ile Ile Ser Tyr Leu Glu Lys Val Ile Pro Leu Ile Ser	
89	195 200 205	
91	gat gca gga ctc cgg tgg cgt ctg aca gat ctg tgc acc aag ata cag	674
92	Asp Ala Gly Leu Arg Trp Arg Leu Thr Asp Leu Cys Thr Lys Ile Gln	
93	210 215 220	
95	cgt ggt acc cac cca gca ata cat gat atc ttg gct ctg aga gag tgt	722
96	Arg Gly Thr His Pro Ala Ile His Asp Ile Leu Ala Leu Arg Glu Cys	
97	225 230 235	
99	ttc tct aac cat cct ctg ggc atg aac caa ctc cag gct ttg cac gtg	770
100	Phe Ser Asn His Pro Leu Gly Met Asn Gln Leu Gln Ala Leu His Val	
101	240 245 250	
103	aaa gcc ttg agc cgg gcc atg ctt ctc aca tct tac ctg cct cct ccc	818
104	Lys Ala Leu Ser Arg Ala Met Leu Leu Thr Ser Tyr Leu Pro Pro Pro	
105	255 260 265 270	
107	ttg ttg aga cat cgt ttg aag act cat aca act gtg att cac caa ctg	866
108	Leu Leu Arg His Arg Leu Lys Thr His Thr Thr Val Ile His Gln Leu	
109	275 280 285	
111	gac aag gct ttg gca aag ctg ggg att ggc cag ctg act gct cag gaa	914
112	Asp Lys Ala Leu Ala Lys Leu Gly Ile Gly Gln Leu Thr Ala Gln Glu	
113	290 295 300	
115	gta aaa tcg gct tgt tat ctc cgt ggc ctg aat tct acg cat att ggt	962
116	Val Lys Ser Ala Cys Tyr Leu Arg Gly Leu Asn Ser Thr His Ile Gly	
117	305 310 315	
119	gaa gat agg tgt cga act tgg ctg gga gaa tgg ctg cag att tcc tgc	1010
120	Glu Asp Arg Cys Arg Thr Trp Leu Gly Glu Trp Leu Gln Ile Ser Cys	
121	320 325 330	
123	agc ctg aaa gaa gct gag ctg tct ctc ttg ctg cac aac gtg gtc ctg	1058
124	Ser Leu Lys Glu Ala Glu Leu Ser Leu Leu Leu His Asn Val Val Leu	
125	335 340 345 350	
127	ctc tcc acc aac tac ctt ggg aca agg cgc tgaatgaacc atggagcgga	1108
128	Leu Ser Thr Asn Tyr Leu Gly Thr Arg Arg	
129	355 360	
131	tggcattgtc ctgcagtcgt atagtatagc agtgcaggaa caaacagcac ttgccagcaa	1168



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,474

DATE: 07/05/2001

TIME: 16:16:43

Input Set : A:\kim.txt

Output Set: N:\CRF3\07032001\I868474.raw

```

133 agtctgtgtg tactgttaag tgtgtgggag gcagagagag gagcaggggc catgggcttc 1228
135 acagcatggc acacctgtgg gaactgcaga cattcctctc acagctagaa ctgaaacaaa 1288
137 ccctcttgct aggggtgggc cgtgtgaggt gtcactcctgt cccctcata attactaata 1348
139 gctggaactg gcagcagcct ctactgggct tttactgtga tgtgttcagt tcatgtccta 1408
141 ggaagtcagc ttttgcccca ggtgggaatc cttatttggc ttaggactga tccacttcca 1468
143 tgttacttac atctgtgggt tttgtgtgtt gctgttagaa aatttttggc tggtgaaaac 1528
145 agcactcctt tggctggagc acttgtgtcc atgcatgtac ttgggtgttt ccctccatcc 1588
147 tttctgatat gacaaaaaat caagttgttt tgtttttggc caccttcact ggcattgggt 1648
149 aaccacttct ttttcaaacc ctctgaacac ctttttctga tgggtaactt gcaggaatat 1708
151 tctattggaa aagataacag gaagtacaag tgcttcttga ccccttctc aatgtttcta 1768
153 gccttcactc tccattgtct tttctgggct gtattacagc cctctgtgga tcttcaactc 1828
155 tgctgcctcc actgtgatgc agcagtccaa ctgtaactga cagtggctgc cttctctggg 1888
157 ccatggatca cacctgtaag gtactaatta ctgccagcc tggggagatc aggagaggtc 1948
159 tgcatagtta gtaagttggg tttagctttt gtgtgtgcat cagtgaacta gagttctgta 2008
161 ataacttatt gtaaattgat gaagcactgt ttttaaacc aagtaaagac tgcttgaaac 2068
163 ctgttgatgg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2118
166 <210> SEQ ID NO: 2
167 <211> LENGTH: 360
168 <212> TYPE: PRT
169 <213> ORGANISM: Homo sapiens
171 <220> FEATURE:
172 <221> NAME/KEY: misc_feature
173 <222> LOCATION: (435)..(494)
174 <223> OTHER INFORMATION: transmembrane domain
176 <400> SEQUENCE: 2
178 Met Ala Leu Ser Arg Val Cys Trp Ala Arg Ser Ala Val Trp Gly Ser
179 1 5 10 15
182 Ala Val Thr Pro Gly His Phe Val Thr Arg Arg Leu Gln Leu Gly Arg
183 20 25 30
186 Ser Gly Leu Ala Trp Gly Ala Pro Arg Ser Ser Lys Leu His Leu Ser
187 35 40 45
190 Pro Lys Ala Asp Val Lys Asn Leu Met Ser Tyr Val Val Thr Lys Thr
191 50 55 60
194 Lys Ala Ile Asn Gly Lys Tyr His Arg Phe Leu Gly Arg His Phe Pro
195 65 70 75 80
198 Arg Phe Tyr Ile Leu Tyr Thr Ile Phe Met Lys Gly Leu Gln Met Leu
199 85 90 95
202 Trp Ala Asp Ala Lys Lys Ala Arg Arg Ile Lys Thr Asn Met Trp Lys
203 100 105 110
206 His Asn Ile Lys Phe His Gln Leu Pro Tyr Arg Glu Met Glu His Leu
207 115 120 125
210 Arg Gln Phe Arg Gln Asp Val Thr Lys Cys Leu Phe Leu Gly Ile Ile
211 130 135 140
214 Ser Ile Pro Pro Phe Ala Asn Tyr Leu Val Phe Leu Leu Met Tyr Leu
215 145 150 155 160
218 Phe Pro Arg Gln Leu Leu Ile Arg His Phe Trp Thr Pro Lys Gln Gln
219 165 170 175
222 Thr Asp Phe Leu Asp Ile Tyr His Ala Phe Arg Lys Gln Ser His Pro
223 180 185 190

```



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,474

DATE: 07/05/2001

TIME: 16:16:43

Input Set : A:\kim.txt

Output Set: N:\CRF3\07032001\I868474.raw

```

226 Glu Ile Ile Ser Tyr Leu Glu Lys Val Ile Pro Leu Ile Ser Asp Ala
227      195      200      205
230 Gly Leu Arg Trp Arg Leu Thr Asp Leu Cys Thr Lys Ile Gln Arg Gly
231      210      215      220
234 Thr His Pro Ala Ile His Asp Ile Leu Ala Leu Arg Glu Cys Phe Ser
235 225      230      235      240
238 Asn His Pro Leu Gly Met Asn Gln Leu Gln Ala Leu His Val Lys Ala
239      245      250      255
242 Leu Ser Arg Ala Met Leu Leu Thr Ser Tyr Leu Pro Pro Pro Leu Leu
243      260      265      270
246 Arg His Arg Leu Lys Thr His Thr Thr Val Ile His Gln Leu Asp Lys
247      275      280      285
250 Ala Leu Ala Lys Leu Gly Ile Gly Gln Leu Thr Ala Gln Glu Val Lys
251      290      295      300
254 Ser Ala Cys Tyr Leu Arg Gly Leu Asn Ser Thr His Ile Gly Glu Asp
255 305      310      315      320
258 Arg Cys Arg Thr Trp Leu Gly Glu Trp Leu Gln Ile Ser Cys Ser Leu
259      325      330      335
262 Lys Glu Ala Glu Leu Ser Leu Leu Leu His Asn Val Val Leu Leu Ser
263      340      345      350
266 Thr Asn Tyr Leu Gly Thr Arg Arg
267      355      360

```

270 &lt;210&gt; SEQ ID NO: 3

271 &lt;211&gt; LENGTH: 18

272 &lt;212&gt; TYPE: DNA

C--&gt; 273 &lt;213&gt; ORGANISM: Artificial

275 &lt;220&gt; FEATURE:

276 &lt;223&gt; OTHER INFORMATION: anti-sense DNA

278 &lt;400&gt; SEQUENCE: 3

279 cctggacatt ttgtcacc

18

282 &lt;210&gt; SEQ ID NO: 4

283 &lt;211&gt; LENGTH: 18

284 &lt;212&gt; TYPE: DNA

C--&gt; 285 &lt;213&gt; ORGANISM: Artificial

287 &lt;220&gt; FEATURE:

288 &lt;223&gt; OTHER INFORMATION: sense DNA

290 &lt;400&gt; SEQUENCE: 4

291 ggtgacaaaa tgtccagg

18

294 &lt;210&gt; SEQ ID NO: 5

295 &lt;211&gt; LENGTH: 18

296 &lt;212&gt; TYPE: DNA

C--&gt; 297 &lt;213&gt; ORGANISM: Artificial

299 &lt;220&gt; FEATURE:

300 &lt;223&gt; OTHER INFORMATION: missense DNA

302 &lt;400&gt; SEQUENCE: 5

303 cgcggatatt tcttcacc

18

306 &lt;210&gt; SEQ ID NO: 6

307 &lt;211&gt; LENGTH: 20

308 &lt;212&gt; TYPE: DNA



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,474

DATE: 07/05/2001

TIME: 16:16:43

Input Set : A:\kim.txt

Output Set: N:\CRF3\07032001\I868474.raw

C--&gt; 309 &lt;213&gt; ORGANISM: Artificial

311 &lt;220&gt; FEATURE:

312 &lt;223&gt; OTHER INFORMATION: forward primer

314 &lt;400&gt; SEQUENCE: 6

315 gggagatgga gcatttgaga

318 &lt;210&gt; SEQ ID NO: 7

319 &lt;211&gt; LENGTH: 20

320 &lt;212&gt; TYPE: DNA

20

C--&gt; 321 &lt;213&gt; ORGANISM: Artificial

323 &lt;220&gt; FEATURE:

324 &lt;223&gt; OTHER INFORMATION: reverse primer

326 &lt;400&gt; SEQUENCE: 7

327 gcttccggaa agcatgatag

20



## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/868,474

DATE: 07/05/2001

TIME: 16:16:44

Input Set : A:\kim.txt

Output Set: N:\CRF3\07032001\I868474.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No  
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:273 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3  
L:285 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4  
L:297 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5  
L:309 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6  
L:321 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7